
Article:
Braunstein, Sam orcid.org/0000-0003-4790-136X and Pirandola, Stefano orcid.org/0000-0001-6165-5615 (2012) Side-Channel-Free Quantum Key Distribution. Physical Review Letters. 130502. ISSN 1079-7114

https://doi.org/10.1103/PhysRevLett.108.130502

Supplementary Material for: Side-channel free quantum key distribution

by Samuel L. Braunstein and Stefano Pirandola

IN DEFENSE OF PRIVATE SPACES

In quantum cryptography unconditional security proofs are derived under the assumption that Alice's and Bob's apparata (private spaces) are completely inaccessible by an eavesdropper who, therefore, can only attack the signal systems which are transmitted through the quantum communication channel connecting the two parties. Under this assumption, secret-key rates and security thresholds are derived in both discrete and continuous variable quantum key distribution.

One potential loophole in the security proofs is related to how a theoretical protocol is actually implemented experimentally. Any redundant information encoded in extra degree of freedom or extra Hilbert space dimensions outside the theoretical prescription can allow for so-called side-channel attacks. By their nature, such attacks may be of classical or quantum degrees of freedom and are insidious because even quantifying their threat appears to involve understanding what have been called unknown unknowns about the vulnerability of the experimental set-up.

Progress has been made on eliminating side channel attacks in the quantum communication channels between private spaces, but this leaves open potential attacks on the private spaces through their quantum communication ports. Let us therefore take a step back and consider private spaces in more details: What goes on in Alice's and Bob's private spaces involves a significant amount of classical information processing; at the very least the key itself will be generated and stored as classical information. Now with virtually any technology we have today classical information is stored, processed and transmitted in a highly redundant fashion (many electrons are used to charge a capacitor to represent a bit value, or many electrons must pass through the base junction of a transistor to effect a logical switching operation, tapping on a keyboard produces sound waves and electromagnetic signals in addition to the 'legitimate' electrical signals in the wires, etc). In principle any of this redundant information may leak out of the private space through a "parasite" channel. An eavesdropper might therefore ignore the quantum communication channel and directly attack Alice's and Bob's apparata by exploiting the presence of parasite channels: this is also a "side-channel attack".

The implicit assumption in quantum cryptography is that we could always improve technology in such a way that Alice's and Bob's private spaces are not affected by the presence of parasite channels, so that the legitimate participants do indeed have access to absolutely private spaces. (For instance, Alice and Bob could simulate the classical information processing on a quantum computer. A hacked operating system on such a machine could be tested for by randomly running subroutines that confirm that coherence is preserved and that no information is copied out to where it can be stored or transmitted by a trojan program — see also Ref. [1].)

However, even if you rely on a perfect isolation technology, there remains a potential chink in this armor, which is the quantum communication port used either to transmit a quantum state out of your private space or to accept a quantum state for detection into it.

If you open a communication port for quantum states to enter or leave you must explicitly deal with side channels which can be probing these links to your private space. Eve can potentially send trojan systems through Alice's and Bob's communication ports and detect their reflection to infer both state preparation and measurement settings. As an example, in the standard BB84 protocol, Eve can irradiate Alice's apparatus by using optical modes at slightly different frequencies. Then, from reflection, Eve can infer the polarization chosen in each round of the protocol. Thanks to this information, Eve can measure each signal system in the correct basis. Another example regards the so-called plug-and-play systems, where trojan systems can be reflected together with signal systems, as discussed in Ref. [2].

Our paper shows how to overcome the problem of the open quantum communication ports, therefore making feasible the notion of absolutely private spaces. Note that this problem is not addressed by current device-independent quantum cryptography, where such attacks on the private space ports are simply considered illegitimate as they violate the strong private space assumption. The key point of our scheme is that detectors are no longer "in line" with the quantum communication port of the private space. For this reason, it is not possible for an external party to probe the port and obtain detector settings or readouts from the processing of parasite systems. In order to explain this key feature in detail, we analyze the problem of the quantum communication ports by comparing standard protocols with our scheme.

In Fig. 1, we depict a general prepare-and-measure protocol, where Alice's variable X is encoded in a quantum state $\rho(X)$ by modulation. Bob's variable Y is the output of a quantum measurement. Here, Eve can attack the quantum communication ports by using two trojan systems e and f. By means of e, Eve can retrieve information about the state preparation $X \to \rho(X)$. By means of f, she can retrieve information about the mea-

surement apparatus of Bob and, therefore, about Y.

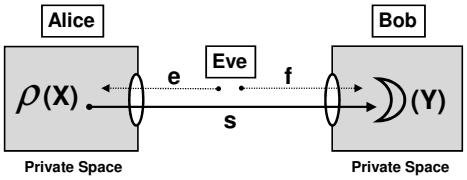

FIG. 1: Port attack in a prepare and measure protocol.

In Fig. 2, we depict a general entanglement-based protocol, where an untrusted party (Eve) distributes entanglement between two parties. This is done by distributing an entangled state $\rho = \rho_{AB}$, where system A is sent to Alice and system B is sent to Bob. Alice and Bob can perform entanglement distillation and measure the output distilled systems to derive two correlated classical variables, X and Y, respectively. In this scenario, Eve can decide not to attack the source ρ but directly the two quantum communication ports of Alice and Bob. Eve can probe these ports by using two trojan systems e and f, which can retrieve information about Alice's and Bob's distilling and detecting apparata. As a result, Eve can infer information about X and Y.

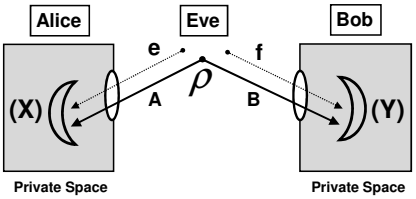

FIG. 2: Port attack in an entanglement-based protocol.

In Fig. 3, we depict our protocol where an untrusted party (Eve) represents an entanglement swapper between Alice and Bob. This is generally done by measuring two *public* systems, A' and B', received from Alice and Bob, processing the outcome of the measurement, and classically communicating the processed data back to Alice and Bob. As a result the two private systems, A and B, become correlated, so that Alice and Bob can extract two correlated classical variables, X and Y, by applying suitable measurements. In particular, if Alice and Bob can access quantum memories, then they can extract a secret key at a rate which is at least equal to the coherent information between A and B. Eve can attempt a side-channel attack against the two ports by sending two trojan systems e and f. In this case, however, the apparata which detect the two private systems A and B are inaccessible to Eve. By exploiting reflections from the ports, Eve can only retrieve information regarding the reduced states $\rho_{A'}$ and $\rho_{B'}$ of the two public systems A' and B'. However, these reduced states contain no useful information about the private system A or B or Alice's or Bob's detector settings or outputs.

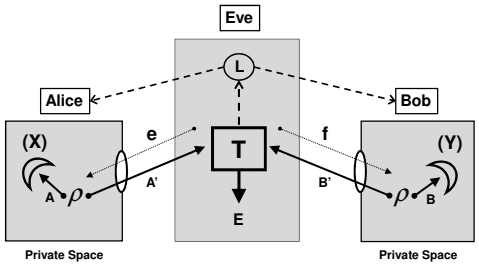

FIG. 3: Port attack in our scheme.

To understand better how the full isolation of the private systems might be achieved, we may consider the procedure depicted in Fig. 4. It is explained for Alice's private space, but steps are identical for Bob.

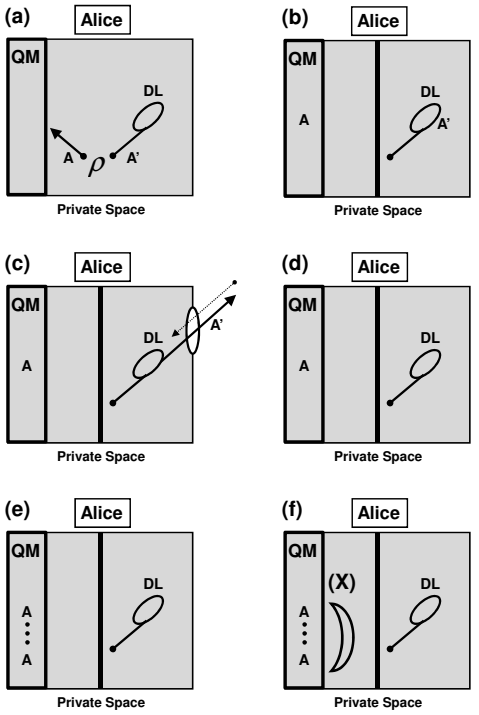

FIG. 4: Possible procedure for the full isolation of the private systems.

In the first step (a), Alice's port is closed and she prepares an entangled state $\rho = \rho_{AA'}$ where system A is directed towards a quantum memory (QM), while system A' is directed towards a delay line (DL). In step (b), once system A is stored in the memory and while system A' is trapped in the delay line, a shutter is used to fully separate the delay line from the rest of Alice's apparatus. Note that a virtual channel between A and A' has been created. In step (c), Alice's quantum communication port is opened and system A' is transmitted to Eve. During this stage, trojan systems may enter the port but

no detector is in line with the port. In step (d), the port is closed with the private system A kept in the memory. The previous steps (a)-(d) are repeated many times, so that Alice collects many private systems in her quantum memory. We therefore reach step (e) of the figure. Finally, once Alice has received all the classical communications, she applies a collective quantum measurement on her quantum memory to retrieve the classical variable X. This measurement can include or be anticipated by an entanglement distillation.

NOTATION AND BASIC FORMULAS

In part of the derivation we adopt the enlarged Hilbert space (EHS) representation, where stochastic classical variables are embedded in quantum systems. Consider a stochastic variable $X = \{x, p(x)\}$ which is encoded into an ensemble of states of some quantum system A, i.e.,

$$\mathcal{E}_A = \{p(x), \rho_A(x)\}. \tag{1}$$

This ensemble may be equivalently represented by the classical-quantum (CQ) state

$$\rho_{\mathbf{X}A} = \sum_x p(x) \left|x\right\rangle \left\langle x\right|_{\mathbf{X}} \otimes \rho_A(x), \tag{2}$$

where the stochastic variable X is embedded into the dummy quantum system $\mathbf{X}$, by using an orthonormal basis $\{|x\rangle\}$ in the Hilbert space $\mathcal{H}_{\mathbf{X}}$ of $\mathbf{X}$. We denote by $\rho_A(x)$ the state of a system A which is conditioned by the value x of a stochastic variable X. The notation $\rho_{A|X}$ refers to the conditional state $\rho_A(x)$ where x is not specified. Clearly, we have

$$\rho_A = \sum_x p(x)\rho_A(x). \tag{3}$$

Given a quantum system A in a state ρ_A, its von Neumann entropy $S(\rho_A)$ is also denoted by $H(A)$. Given a quantum system $\mathbf{X}$, embedding the stochastic variable X, its quantum entropy $H(\mathbf{X})$ is just the Shannon entropy $H(X)$. Given two quantum systems, A and B, we denote by $I(A:B)$ their quantum mutual information. This is defined by

$$I(A:B) = H(B) - H(B|A), \tag{4}$$

where

$$H(B|A) = H(AB) - H(A), \tag{5}$$

is the conditional quantum entropy. Note that $H(B|A)$ can be negative and it is related to the coherent information by the relation

$$I(A\rangle B) = -H(B|A). \tag{6}$$

For $A = \mathbf{X}$, the quantum mutual information $I(A:\mathbf{X})$, which is computed over the CQ-state of Eq. (2), corresponds to the Holevo information $I(A:X)$, computed over the ensemble of Eq. (1). For $A = \mathbf{X}$ and $B = \mathbf{Y}$, embedding two stochastic variables X and Y, $I(\mathbf{X}:\mathbf{Y})$ is just the classical mutual information $I(X:Y)$. For three quantum systems A, B, and C, we can consider the conditional quantum mutual information

$$I(A:B|C) = H(AC) + H(BC) - H(ABC) - H(C), \tag{7}$$

which is ≥ 0 as a consequence of the strong subadditivity of the von Neumann entropy. For a classically correlated system $C = \mathbf{X}$, we have a probabilistic average over mutual informations, i.e.,

$$I(A:B|\mathbf{X}) = I(A:B|X) \equiv \sum_x p(x)\ I(A:B|X=x). \tag{8}$$

List of other useful elements:

- Given a tripartite quantum system ABC, we can use the "chain rule"

$$I(A:BC) = I(A:B) + I(A:C|B). \tag{9}$$

- Invariance of the Holevo information under addition of classical channels, i.e., for a classical channel

$$p(y|x) : X \to Y, \tag{10}$$

we have

$$I(A:X) = I(A:XY). \tag{11}$$

- Given a Markov chain $X \to Y \to Z$, the classical mutual information decreases under conditioning [3], i.e.,

$$I(X:Y|Z) \leq I(X:Y). \tag{12}$$

Notice that, for three general stochastic variables, we have $I(X:Y|Z) \gtrless I(X:Y)$, so that the so-called "interaction information"

$$I(X:Y:Z) \equiv I(X:Y|Z) - I(X:Y), \tag{13}$$

can be positive, negative or zero.

- Data processing inequality. For a Markov chain $X \to Y \to Z$, we have

$$H(X) \geq I(X:Y) \geq I(X:Z). \tag{14}$$

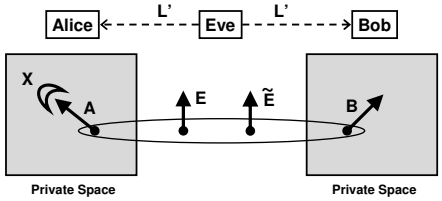

FIG. 5: Purification. Conditional state $\Phi_{ABE\tilde{E}|L'}$ projected onto $\Phi_{BE\tilde{E}|XL'}$.

PROOF OF THE THEOREM

Let us purify the mixed state $\rho_{ABE|L'}$ into the pure state $\Phi_{ABE\tilde{E}|L'} = |\Phi\rangle\langle\Phi|_{ABE\tilde{E}|L'}$ by introducing an ancillary system $\tilde{E}$ which is assumed to be in Eve's hands (so that Eve's global system consists of $E\tilde{E}$). This scenario is depicted in Fig. 5.

Thus, for the total state $\rho_{ABE|L'}$, we have

$$\rho_{ABE}(l') = \mathrm{Tr}_{\tilde{E}}\left[\Phi_{ABE\tilde{E}}(l')\right]. \tag{15}$$

For the conditional state $\rho_{BE|XL'}$, generated by the measurement, we can write

$$\begin{aligned}\rho_{BE}(x,l') &= \frac{1}{p(x|l')}\mathrm{Tr}_A\left[\hat{A}(x)\rho_{ABE}(l')\hat{A}(x)^\dagger\right]\\ &= \frac{1}{p(x|l')}\mathrm{Tr}_{A\tilde{E}}\left[\hat{A}(x)\Phi_{ABE\tilde{E}}(l')\hat{A}(x)^\dagger\right]\\ &= \mathrm{Tr}_{\tilde{E}}\left[\Phi_{BE\tilde{E}}(x,l')\right],\end{aligned} \tag{16}$$

where

$$\Phi_{BE\tilde{E}}(x,l') \equiv \frac{1}{p(x|l')}\mathrm{Tr}_A\left[\hat{A}(x)\Phi_{ABE\tilde{E}}(l')\hat{A}(x)^\dagger\right], \tag{17}$$

represents the conditional state $\Phi_{BE\tilde{E}|XL'}$ which is generated by the measurement in the purified scenario. Clearly if we discard X, we get the reduced state

$$\Phi_{BE\tilde{E}|L'} \equiv \left\langle \Phi_{BE\tilde{E}|XL'}\right\rangle_X = \mathrm{Tr}_A\left[\Phi_{ABE\tilde{E}|L'}\right]. \tag{18}$$

Because of Eq. (16), the conditional state $\Phi_{BE\tilde{E}|XL'}$ can be used to compute R' via

$$\begin{aligned}R' &\equiv I(X:B|L')_\rho - I(X:E|L')_\rho\\ &= I(X:B|L')_\Phi - I(X:E|L')_\Phi,\end{aligned} \tag{19}$$

where $\rho = \rho_{BE|XL'}$ and $\Phi = \Phi_{BE\tilde{E}|XL'}$ (the computation is exactly the same up to a trace over $\tilde{E}$). In the EHS representation, the conditional state $\Phi_{BE\tilde{E}|XL'}$ becomes

$$\Psi_{\mathbf{XL}'BE\tilde{E}} = \sum_{x,l'} p(x,l')\,|x\rangle\langle x|_{\mathbf{X}} \otimes |l'\rangle\langle l'|_{\mathbf{L}'} \otimes \Phi_{BE\tilde{E}}(x,l'). \tag{20}$$

Thus, we can also set

$$R' = I(\mathbf{X}:B|\mathbf{L}')_\Psi - I(\mathbf{X}:E|\mathbf{L}')_\Psi, \tag{21}$$

where $\Psi = \Psi_{\mathbf{XL}'BE\tilde{E}}$. From the chain rule we have

$$\begin{aligned}I(\mathbf{X}:E\tilde{E}|\mathbf{L}')_\Psi &= I(\mathbf{X}:E|\mathbf{L}')_\Psi + I(\mathbf{X}:\tilde{E}|E\mathbf{L}')_\Psi\\ &= I(\mathbf{X}:E|\mathbf{L}')_\Psi + \gamma,\end{aligned} \tag{22}$$

where $\gamma \equiv I(\mathbf{X}:\tilde{E}|E\mathbf{L}')_\Psi \geq 0$ is the information contribution due to the purification [4]. In other words, the (conditional) Holevo information can only increase with the purification, i.e.,

$$I(X:E\tilde{E}|L') = I(X:E|L') + \gamma \geq I(X:E|L'). \tag{23}$$

As a consequence, we have $R' = R'' + \gamma$, where

$$R'' \equiv I(X:B|L')_\Phi - I(X:E\tilde{E}|L')_\Phi. \tag{24}$$

In terms of conditional entropies, we have

$$\begin{aligned}R'' =& H(B|L')_\Phi - H(B|XL')_\Phi\\ &- [H(E\tilde{E}|L')_\Phi - H(E\tilde{E}|XL')_\Phi].\end{aligned} \tag{25}$$

Here $H(E\tilde{E}|L')$ is computed over $\Phi = \Phi_{BE\tilde{E}|XL'}$ discarding X and B, i.e., over the reduced state

$$\Phi_{EE|L'} = \mathrm{Tr}_{AB}\left[\Phi_{ABE\tilde{E}|L'}\right]. \tag{26}$$

Now since $\Phi_{ABE\tilde{E}|L'}$ is pure, we have $H(E\tilde{E}|L') = H(AB|L')$, where $H(AB|L')$ can be computed over $\rho_{AB|L'} = \mathrm{Tr}_{E\tilde{E}}[\Phi_{ABE\tilde{E}|L'}]$. Clearly, also $H(B|L')_\Phi$ can be computed over $\rho_{AB|L'}$. As a consequence we can recognize in Eq. (25) the conditional coherent information

$$I(A\rangle B|L') = H(B|L') - H(AB|L'),$$

associated with Alice and Bob's conditional state $\rho_{AB|L'}$. Thus, we can set

$$R'' = I(A\rangle B|L') + [H(E\tilde{E}|XL')_\Phi - H(B|XL')_\Phi]. \tag{27}$$

Here, we can assume that Alice's measurement is a rank one POVM. As a result, $\Phi = \Phi_{BE\tilde{E}|XL'}$ is also a pure state, and we can set $H(E\tilde{E}|XL')_\Phi = H(B|XL')_\Phi$, so that $R'' = I(A\rangle B|L')$. Finally, we can write

$$\begin{aligned}R^* &= R'' + \gamma + \Delta\\ &= I(A\rangle B|L') + \gamma + \Delta\\ &\geq I(A\rangle B|L') + \Delta,\end{aligned} \tag{28}$$

where we have used $\gamma \geq 0$ from its definition.

[1] S. Barz *et al.*, Science **335**, 303 (2012).
[2] N. Gisin *et al.*, Rev. Mod. Phys. **74**, 145-195 (2002)
[3] T. M. Cover and J. A. Thomas, (John Wiley and Sons, Hoboken, New Jersey, 2006) p. 35.
[4] Note that the EHS representation has been mainly introduced to give the correct interpretation to Eq. (), where a quantum system E conditions a classical variable X thanks to the embedding in a quantum system $\mathbf{X}$.